Original Article

Conversational Pharmacovigilance That Preserves the Patient's Narrative

Olivia Bennett[1]*, Harry Collins[2], Jack Foster[1], Ethan Wright[3]

[1]Department of Conversational AI and Pharmacovigilance, Faculty of Pharmacy, University of Nottingham, Nottingham, United Kingdom.
[2]Department of Patient Narrative in Drug Safety, Faculty of Pharmacy, University of Southampton, Southampton, United Kingdom.
[3]Department of AI and Patient Voice in Pharmacy, Faculty of Pharmacy, Newcastle University, Newcastle, United Kingdom.

Abstract

Structured adverse-event reporting supports coding, aggregation, and regulatory review, but it can compress the patient's account into fields that omit chronology, uncertainty, functional consequences, perceived causality, and the language through which harm is understood. This article proposes a non-empirical conversational pharmacovigilance architecture designed to preserve patient narrative while supporting structured extraction, triage, and signal work. The architecture uses a dual-representation contract: an append-preserving narrative ledger remains the source record, while a separate structured safety object contains derived medication, event, timing, seriousness, outcome, action, and uncertainty fields linked to supporting narrative spans. A bounded clarification controller may request missing safety-relevant information without silently rewriting previous statements. The preserved narrative and structured derivative are combined in a paired triage packet and passed through explicit human review gates before case assessment or signal work. Governance spans consent, privacy, accessibility, patient correction, provenance, auditability, subgroup evaluation, incident response, change control, and withdrawal authority. Evaluation must distinguish narrative completeness and fidelity from coding accuracy, triage safety, workflow consequence, patient experience, equity, and organizational readiness. The architecture does not determine causality, replace professional assessment, provide medication advice, validate safety signals, or establish regulatory compliance. Its original contribution is a provenance-preserving relationship between patient dialogue and pharmacovigilance operations that treats structured data as a useful derivative rather than a superior replacement for the patient's account. Empirical validation is required across languages, populations, reporting channels, and medication-safety workflows.

Keywords: Digital pharmacy, Artificial intelligence, Pharmacy practice, Medication safety, Human–AI collaboration, Clinical decision support

Introduction

Spontaneous reporting systems require standardization. Drug names, suspected reactions, timing, seriousness, outcomes, and reporter characteristics must be represented consistently enough to permit case processing, retrieval, aggregation, and signal assessment. Yet standardization changes the form of evidence. Patient reports can contain clinically useful information comparable with professional reports in several report-quality dimensions, challenging assumptions that patient-authored accounts are inherently inferior [1].

The problem is therefore not whether structure is necessary, but whether structure becomes the only representation retained.

A patient narrative may communicate how an event emerged, what changed in daily life, why the patient associated it with treatment, what actions were taken, and what remained uncertain. Qualitative analysis of Yellow Card reports showed that patients described recognition, experienced impact, and management of suspected adverse reactions in ways that extend beyond reaction labels [2]. When such material is condensed into discrete fields, the result may remain administratively complete while becoming semantically thinner. "Narrative meaning" in this article denotes the patient's wording, sequence, uncertainty, attribution, experienced consequences, priorities, and revisions—not an assumption that every narrative is factually complete or causally correct.

The reporting channel also shapes what becomes visible. Patient-controlled online reporting has surfaced sensitive

Address for correspondence: Olivia Bennett, Department of Conversational AI and Pharmacovigilance, Faculty of Pharmacy, University of Nottingham, Nottingham, United Kingdom.
olivia.bennett@nottingham.ac.uk
Received: 26 January 2026; **Accepted:** 30 March 2026

How to cite this article: Bennett O, Collins H, Foster J, Wright E. Conversational Pharmacovigilance That Preserves the Patient's Narrative. Arch Pharm Pract. 2026;17(2):4-12. https://doi.org/10.51847/HSMLaWFbcN

adverse experiences, including sexual events that may be difficult to disclose through conventional encounters [3]. Conversely, conversational systems may introduce their own distortions through leading questions, excessive prompts, anthropomorphic reassurance, inaccessible language, or pressure to complete predefined fields. Narrative preservation is therefore not achieved merely by adding a chat interface.

Underreporting further complicates interpretation. Patient reporting is affected by awareness, beliefs, burden, perceived legitimacy, and practical access, so absence of a report cannot be treated as absence of an adverse experience [4]. A conversational pathway may reduce some barriers while creating others. This article consequently proposes an architecture rather than a validated reporting tool. Its central claim is that pharmacovigilance can support structured work without allowing the structured derivative to overwrite its narrative source. The proposal separates capture from interpretation, extraction from triage, and triage from professional case and signal decisions.

Patient Narratives as Pharmacovigilance Evidence

Patient narratives are pharmacovigilance evidence because they can contribute information about temporal emergence, lived consequences, management behavior, uncertainty, and patient-defined importance. Their value is complementary rather than self-validating. In global reporting data, patient reports sometimes entered the system earlier than professional reports, indicating a potentially relevant temporal contribution [5]. This does not mean that every patient report accelerates detection or that early reporting confirms a causal signal.

Patient-originated reports can also introduce concerns or experiential dimensions not adequately represented in professional reporting. Collaborative signal work using VigiBase identified safety concerns originating from patient reports and illustrated how patient interpretation can enrich formal review [6]. Such contributions should remain distinguishable from regulatory conclusions: a reported concern is evidence for assessment, not proof of causation.

Narrative availability is itself shaped by participation. Patients and health professionals report for different, context-dependent reasons, including perceived seriousness, desire to prevent harm, expectations of feedback, and confidence that reporting is worthwhile [7]. A conversational system that ignores these determinants may collect technically complete fields while failing to support meaningful participation. Patient-centered pharmacovigilance therefore also concerns understandable communication, involvement, and shared responsibility for medication safety [8].

Taken together, patient reports can contribute temporally relevant signal material, experiential interpretation, and context-sensitive participation evidence [5–7]. The proposed architecture treats these as three related but non-equivalent dimensions. Temporal contribution concerns when information becomes available; experiential contribution concerns what the patient uniquely describes; participatory contribution concerns whether the reporting process permits that evidence to appear. None independently determines causality, seriousness, preventability, or required action.

Table 1 organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for definitions, components, relationships, and intended uses for the article Conversational Pharmacovigilance That Preserves the Patient's Narrative.

Table 1. Definitions, components, relationships, and intended uses for the article Conversational Pharmacovigilance That Preserves the Patient's Narrative.

Component or scientific dimension	Problem addressed	Inputs or determinants	Proposed mechanism or relationship	Expected contribution	Evidence required	Failure or uncertainty risk	Boundary statement
Patient narrative	Structured fields may omit lived meaning.	Patient wording, chronology, impact, and uncertainty	**Proposed:** retain the source account as the primary representation.	Preserves context for review.	Narrative–structured comparison [2]	Incomplete, inconsistent, or causally mistaken account	Narrative richness does not establish causality.
Comparative report quality	Patient evidence may be discounted.	Clinical detail and report completeness	Use reporter type as context, not as an evidence hierarchy.	Supports inclusion of patient-authored evidence.	Comparative quality assessment [1]	Quality varies across fields and cases.	Comparable information is not universal equivalence.
Sensitive disclosure	Conventional channels may suppress difficult experiences.	Privacy, stigma, and channel control	**Proposed:** offer private, interruptible reporting routes.	May broaden visible safety experience.	Channel-specific disclosure evidence [3]	Digital exclusion or new privacy risks	Access does not prove completeness or validity.
Underreporting	Unreported experience remains invisible.	Awareness, beliefs, effort, and feedback expectations	**Proposed:** reduce avoidable burden without coercion.	May support participation.	Patient underreporting evidence [4]	Abandonment, selection bias, or false reassurance	Non-reporting is not evidence of non-occurrence.

Temporal contribution	Emerging concerns may reach systems at different times.	Reporter recognition and reporting delay	Preserve patient timing and uncertainty explicitly.	Supports time-sensitive review.	Time-to-reporting evidence [5]	Recall error and delayed recognition	Earlier reporting does not validate a signal.
Experiential contribution	Codes may omit functional and personal consequences.	Daily-life impact, management, and patient priorities	**Proposed:** link structured fields to source narrative spans.	Retains clinically interpretable context.	Patient-signal evidence [6]	Overinterpretation of subjective attribution	Experience is evidence for assessment, not proof.
Participatory contribution	Reporting behavior is not neutral.	Motivation, trust, burden, and anticipated response	**Proposed:** disclose purpose and provide meaningful acknowledgement.	Supports legitimate participation.	Multicountry motives evidence [7]	Unequal participation and expectation mismatch	Engagement does not establish representativeness.
Patient-centered governance	Technical collection may exclude patient agency.	Choice, communication, correction, and involvement	**Proposed:** provide correction, annotation, and withdrawal pathways.	Supports dignity and accountability.	Patient-centered priorities [8]	Tokenistic involvement or unusable controls	Patient-centeredness does not transfer professional authority.

Proposed Conversational Pharmacovigilance Architecture

Artificial intelligence can support pharmacovigilance only when its role is defined within the broader medicine-safety lifecycle. Automation may assist intake, coding, retrieval, and prioritization, but it does not independently perform causal assessment or accountable signal decisions [9]. The proposed architecture therefore centers not on a model but on a controlled relationship among patient dialogue, derived data, human review, and governance.

Conversational and telehealth channels may extend access to patient-generated safety information while introducing documentation, privacy, interoperability, and workflow constraints [10]. Mobile reporting similarly offers additional access routes, yet availability does not demonstrate better report quality, safer triage, or improved pharmacovigilance outcomes [11]. The architecture consequently remains channel-neutral: it may be implemented through text, voice, assisted reporting, or hybrid professional–patient interaction, but each channel requires separate validation.

The central mechanism is a dual-representation contract. First, an append-preserving narrative ledger stores each conversational turn with its speaker, time, language, correction status, and provenance. Second, a structured safety object stores derived medication, event, timing, seriousness indicators, experienced impact, action, outcome, missingness, and uncertainty. Structured fields must link to supporting source spans. Clarifications are appended and may resolve ambiguity, but they may not silently replace earlier wording or conceal contradictions.

Seven proposed components organize the architecture: conversational access and role disclosure; the narrative ledger; a bounded clarification controller; structured extraction and provenance; a safety and triage router; a human signal-work interface; and a governance and assurance spine. Human gates separate capture from interpretation, extraction from triage, triage from case assessment, and case assessment from signal action. Validation, oversight, and quality controls are required before operational AI use in pharmacovigilance [12].

Figure 1 presents the original conceptual synthesis for conversational pharmacovigilance architecture, showing how the article's principal components, evidence relationships, uncertainties, and decision boundaries are connected.

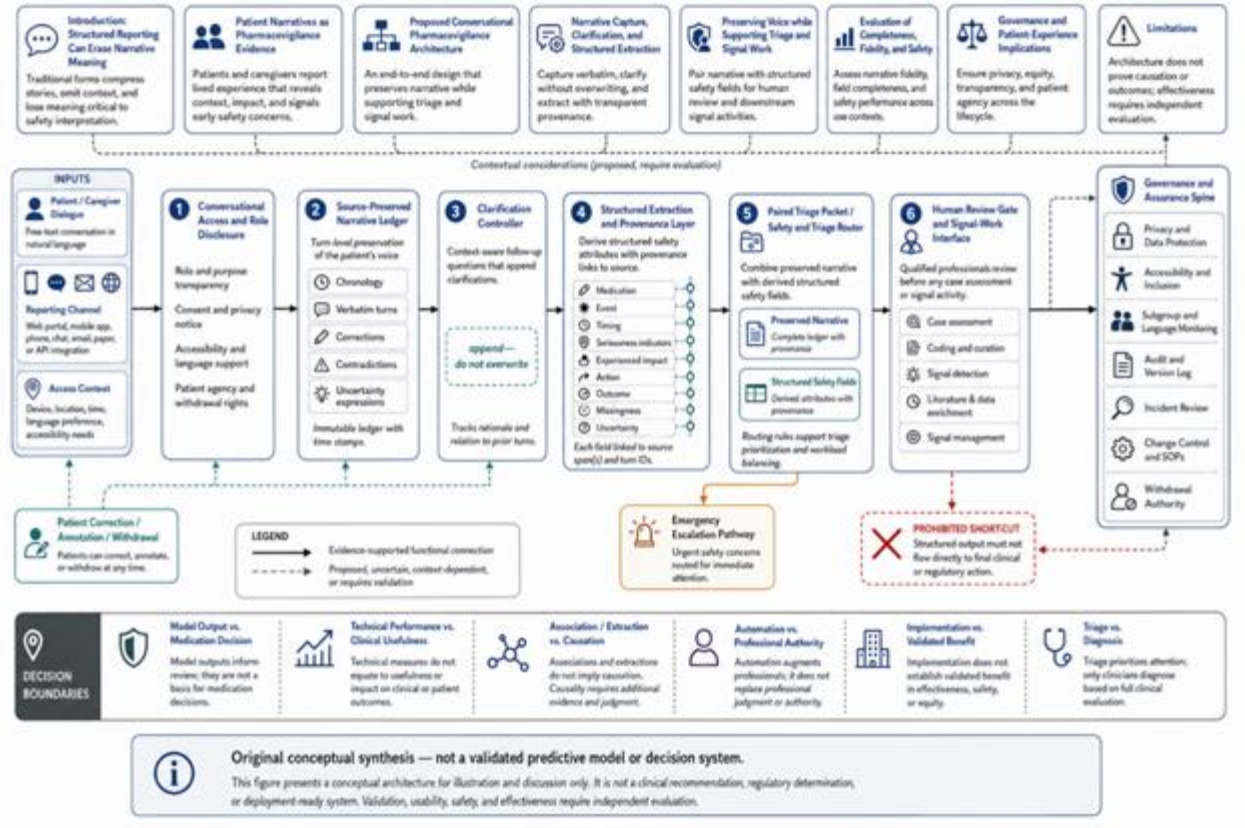

Figure 1. Conversational Pharmacovigilance Architecture

The architecture permits reversible movement. A patient may correct, annotate, pause, or withdraw where legally and operationally permissible; the system retains amendment history rather than fabricating a single frictionless account. An urgent-risk pathway may escalate information to a qualified professional, but the conversational system must not offer unsupported reassurance, diagnose the event, or modify therapy. Signal teams receive the preserved narrative and structured derivative together, not a decontextualized code set.

Table 2 organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for evaluation criteria, evidence requirements, and failure modes for the article Conversational Pharmacovigilance That Preserves the Patient's Narrative.

Table 2. Evaluation criteria, evidence requirements, and failure modes for the article Conversational Pharmacovigilance That Preserves the Patient's Narrative.

Architecture layer or process stage	Core function	Information flow	Interaction with other components	Evaluation criterion	Potential failure mode	Human or organizational responsibility	Readiness boundary
Access and role disclosure	Establish purpose, limits, consent, and accessibility.	Patient context → dialogue channel	Precedes all capture.	Comprehension, accessibility, and voluntary use	Deceptive role, exclusion, or coerced completion	Service owner and patient-experience lead	No use where system role or emergency limits are unclear.
Narrative ledger	Preserve source turns and amendments.	Dialogue → source record	Supplies extraction and human review.	Completeness, chronology, and correction traceability	Silent rewriting or lost contradictions	Data custodian and medication-safety lead	No downstream reliance without recoverable source text.
Clarification controller	Request missing safety-relevant detail.	Missingness signal ↔ patient response	Reads the ledger and appends responses.	Nonleading, proportionate, and interruptible questioning	Coaching, distress, or abandonment	Clinical safety owner and human-factors team	No autonomous clarification outside the validated scope.
Structured extraction	Derive coded safety fields.	Narrative spans → structured object	Feeds triage with provenance.	Source-linked accuracy and	Unsupported inference or false certainty	Model steward and pharmacovigilance expert	No untraceable field may support

				uncertainty handling			consequential use.
Triage router	Prioritize review without adjudicating causality.	Narrative + fields + uncertainty → review queue	Precedes human assessment.	Urgent-case recognition and acceptable escalation burden	Missed urgency, false reassurance, or alarm overload	Qualified safety reviewer and service operator	No autonomous closure, diagnosis, or treatment advice.
Human signal-work interface	Support verification and broader evidence assessment.	Paired packet → professional review	Integrates other pharmacovigilance evidence.	Verifiability, workload, and appropriate reliance	Automation bias or narrative neglect	Pharmacovigilance organization	No signal action based solely on model output.
Governance and assurance spine	Control privacy, equity, monitoring, incidents, and updates.	Audit findings → review, restriction, change, or withdrawal	Spans every layer.	Accountable ownership and controlled lifecycle [12]	Responsibility gaps or undocumented change	Accountable organization and governance board	No operation without monitoring and withdrawal authority.
End-to-end architecture	Maintain separation among evidence, derivation, and authority.	Patient dialogue → paired evidence → human decision	Coordinates all layers.	Context-specific validation and workflow fit [9]	Technical success mistaken for validated benefit	Institutional decision-maker	Conceptual coherence is not deployment readiness.

All architecture mechanisms and readiness boundaries are proposed unless explicitly identified as evidence-derived.

The architecture is intentionally non-autonomous. It may organize evidence and support professional attention, but it does not establish causality, validate a signal, determine treatment, satisfy regulatory requirements, or justify use beyond the population, language, channel, and workflow in which it has been evaluated.

Narrative Capture, Clarification, and Structured Extraction

Conversational capture should be understood as a sequence of functions rather than a single chatbot capability. Healthcare conversational agents vary substantially in purpose, modality, dialogue control, and evidential maturity [13]. A pharmacovigilance system must therefore state whether it is collecting information, clarifying ambiguity, providing navigation, or supporting professional review. These functions should not be blended in ways that make advice appear to be neutral data collection.

Personalization can alter question order, language, reminders, and response style [14]. In this architecture, personalization is permitted only when it improves accessibility or resolves a defined information need. It must remain visible, auditable, and bounded. The system should not infer emotional states, change risk messages to maximize completion, or steer patients toward preferred causal interpretations. Adaptation that cannot be explained at the interaction level should not be used for safety-critical clarification.

A clarification controller is proposed to identify missing or conflicting elements such as medication identity, temporal sequence, action taken, outcome, or possible urgency. The distinction between clarification and interrogation is essential. Conceptual analyses of health agents show that information gathering, support, education, and decision functions are different conversational roles [15]. Questions should therefore be minimal, nonleading, interruptible, and sensitive to distress. "I am not sure," "I do not remember," and contradictory responses must remain valid data states rather than triggers for forced completion.

Structured extraction begins only after source preservation. Automated coding of patient adverse-reaction reports can support expert work when validated against real pharmacovigilance data [16]. The proposed structured safety object may contain coded medication and event concepts, dates or intervals, seriousness indicators, experienced impact, management actions, outcomes, reporter identity, source-span links, and uncertainty states. Every field must be classified as explicit, inferred, conflicting, missing, or unsupported.

The narrative ledger remains authoritative for what the patient said; the structured object remains authoritative only for what the system derived. A summary must be labelled as derived, and every consequential field must be traceable to the exact supporting turn. Corrections are appended with history preserved. Conflicts are surfaced to the reviewer rather than reconciled invisibly. These rules permit efficient coding and triage while protecting against a category error: assuming that a clean structured record is a more truthful account than the uncertain narrative from which it was produced.

Preserving Voice while Supporting Triage and Signal Work

Transforming narrative into standardized terminology may improve retrieval and comparison, but coding remains a derived representation. Artificial intelligence has been applied to coding patient-reported adverse reactions, demonstrating that unstructured accounts can support structured pharmacovigilance processing [17]. Such conversion does not establish that the coded term preserves

chronology, emphasis, functional impact, uncertainty, or the patient's reason for considering the experience important.

Free-text detection also depends on the language, documentation practices, and context in which a method is developed. Adverse reactions have been identified from electronic health-record notes, but performance and interpretation remain conditioned by local documentation patterns [18]. Patient dialogue differs from clinical notes because it may be less standardized, more iterative, multilingual, emotionally expressed, or uncertain. Validation based on one text source should therefore not be transferred automatically to another.

Signal work introduces a further boundary. Electronic health-record methods for identifying possible drug-safety signals are heterogeneous in design, data quality, and evidential purpose [19]. A candidate generated from narrative extraction is consequently an input to assessment rather than a confirmed signal. Reviews of natural-language processing and machine-learning approaches for adverse-event detection likewise indicate methodological heterogeneity and an evidence base that does not justify unsupported clinical autonomy [20].

Automated coding, free-text detection, and electronic-record signal methods can support retrieval and prioritization, but each remains a context-dependent input to expert assessment [17–19]. The proposed architecture operationalizes this boundary through a paired triage packet containing the preserved narrative, derived fields, source-span provenance, contradictions, missingness, uncertainty, and urgency indicators. The reviewer can inspect why a case was prioritized and return to the patient's exact wording.

Triage is defined as reversible prioritization, not diagnosis, causality assessment, case closure, or signal confirmation. A system may flag possible urgency or incomplete information, but a qualified professional must verify consequential interpretations. The architecture prohibits a direct pathway from structured fields to medication advice, regulatory action, or final signal determination. Preserving voice therefore does not mean avoiding structure; it means ensuring that structure remains traceable, contestable, and subordinate to accountable interpretation.

Evaluation of Completeness, Fidelity, and Safety

Evaluation must distinguish whether the conversation collected required information from whether it represented the patient faithfully. Existing evaluations of healthcare conversational agents span functionality, information quality, safety, user experience, outcomes, uptake, costs, and user characteristics [21]. A pharmacovigilance architecture should add narrative-specific domains: preservation of sequence, attribution, uncertainty, patient-prioritized impact, corrections, and contradictions.

Early-stage evaluation should examine intended use, workflow integration, human factors, system errors, and the clinical environment before progression to larger evaluations [22]. Retrospective coding performance alone cannot reveal whether patients abandon the conversation, whether clarification questions distort answers, whether reviewers over-rely on structured summaries, or whether urgent reports are safely escalated.

Transparent assessment must document system inputs, human–AI interaction, error handling, and version changes [23]. Model-oriented reporting should also address data provenance, performance, fairness, applicability, and limitations [24]. Reporting quality, however, does not itself establish clinical usefulness or safety.

A staged validation program must integrate conversational-function evaluation, early workflow and human-factors assessment, and transparent reporting of AI intervention behavior [21–23]. Proposed evaluation domains are:

- Completeness: whether safety-relevant details and patient-prioritized consequences are captured.
- Fidelity: whether source meaning, chronology, uncertainty, and attribution are retained.
- Extraction validity: whether derived fields are accurate, source-linked, and appropriately uncertain.
- Triage safety: whether urgent cases are recognized without unsafe reassurance or excessive false escalation.
- Workflow consequence: whether reviewers can verify outputs without excessive burden or automation bias.
- Patient experience: whether participation is understandable, voluntary, dignified, and correctable.
- Equity: whether failures and burdens differ across languages, literacy levels, disabilities, communication styles, or underserved groups.
- Governance readiness: whether monitoring, incident response, version control, and withdrawal authority are operational.

No universal readiness score or unsupported numerical threshold is proposed. Criteria and transition limits must be prespecified for each intended population, language, reporting channel, and workflow.

Governance and Patient-Experience Implications

Responsible health AI requires governance from problem formulation through deployment, monitoring, and accountability [25]. For conversational pharmacovigilance, governance must identify who owns the narrative record, who validates extraction, who reviews urgent cases, who investigates incidents, and who may restrict or withdraw the system.

Privacy, transparency, bias, accountability, and human agency are distinct ethical obligations rather than interchangeable indicators of responsible design [26]. Patients should be told when they are interacting with an

automated system, what it can and cannot do, how information will be used, and when a professional becomes responsible. Correction and withdrawal mechanisms should be usable rather than merely disclosed.

Patient acceptance of healthcare AI is conditional on safety, choice, data security, bias, costs, and retained human involvement [27]. Aggregate technical performance may also conceal differential missed-event risks for underserved populations [28]. Evaluation must therefore examine who is not reporting, whose language is poorly represented, whose narratives are disproportionately compressed, and whose urgent experiences are under-triaged.

Figure 2 presents the original conceptual synthesis for narrative preservation from patient dialogue to triage and signal work, showing how the article's principal components, evidence relationships, uncertainties, and decision boundaries are connected.

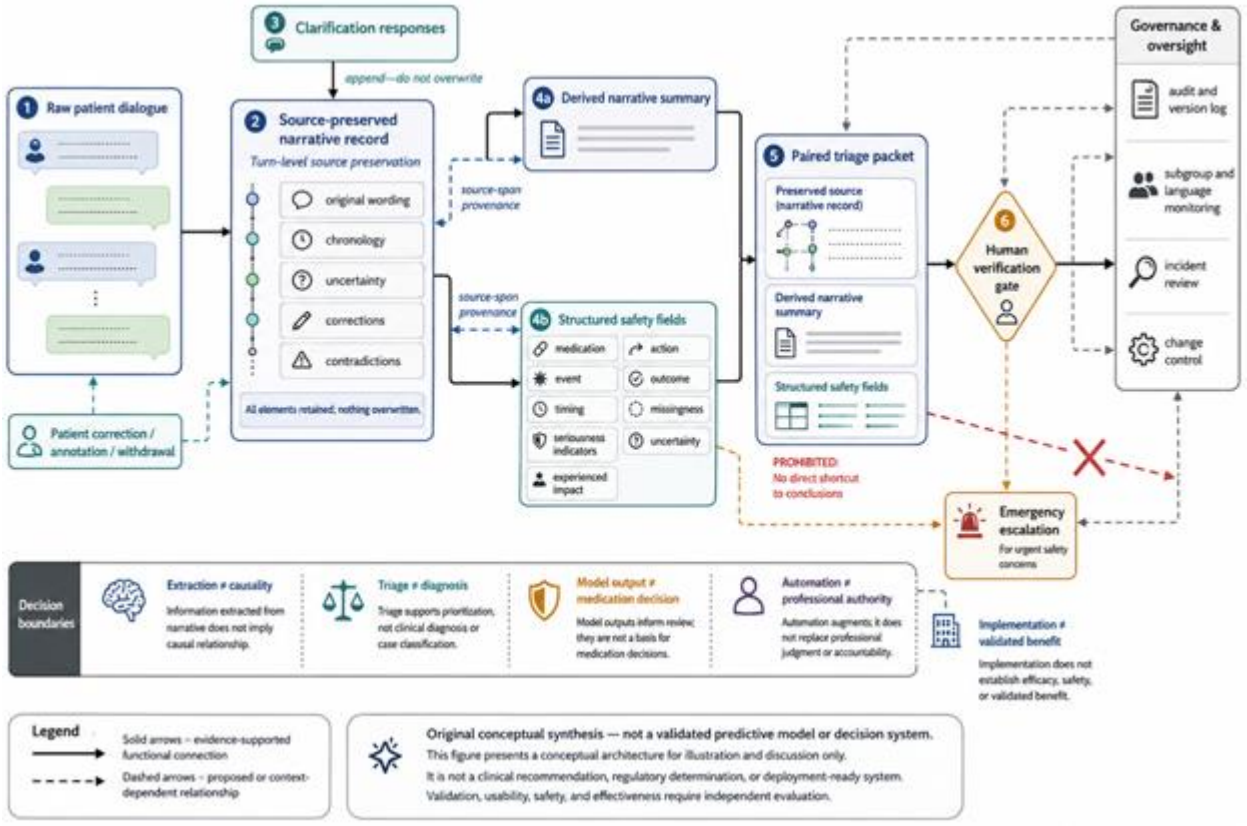

Figure 2. Narrative Preservation from Patient Dialogue to Triage and Signal Work

The figure's correction loop represents narrative custodianship: the organization retains responsibility for source integrity while the patient retains a meaningful ability to contest derived representations. Governance must also preserve model versions, prompt policies, transformation histories, reviewer actions, and subgroup performance. These mechanisms are proposed safeguards; they do not independently establish legal compliance, regulatory acceptance, equitable performance, patient trust, or organizational readiness.

Limitations

Clinical impact cannot be inferred from model performance without external validation, workflow fit, implementation capacity, and adoption [29]. Explanations generated for model outputs should not be treated as proof of correctness, causality, or safety [30]. Medical AI systems may also be vulnerable to adversarial manipulation and other security failures [31]. Performance can change across sites, populations, languages, and time because of dataset shift [32].

The architecture may additionally fail through inaccessible dialogue, conversational coercion, mistaken medication identity, incomplete chronology, culturally mismatched clarification, unrecognized distress, or excessive reviewer workload. Preservation may conflict with data-minimization requirements, and patient correction may be constrained after information enters formal safety records. The proposal therefore requires context-specific legal, technical, clinical, human-factors, equity, and organizational evaluation. Conceptual plausibility does not establish generalizability or decision readiness.

Conclusion

Conversational pharmacovigilance should not force a choice between unstructured patient voice and operationally useful safety data. This article proposes a dual-representation architecture in which the original dialogue remains available through an append-preserving narrative ledger, while structured fields are derived separately, linked to source spans, and accompanied by uncertainty. Clarification is bounded; triage is reversible; and professional review separates model output from consequential medication-safety decisions.

The architecture's scientific contribution is the explicit treatment of narrative fidelity as distinct from completeness, coding accuracy, and technical performance. Its professional contribution is a paired interface that allows pharmacists and pharmacovigilance reviewers to verify structured information against patient wording. Its organizational contribution is a governance spine covering patient agency, provenance, accountability, monitoring, incident response, equity assessment, and controlled change.

The proposal remains non-empirical. It does not establish causality, validate a safety signal, recommend treatment, replace professional authority, satisfy regulatory requirements, or demonstrate benefit. Validation must determine whether the architecture preserves meaning, improves usable completeness, supports safe triage, avoids disproportionate burden or exclusion, and functions acceptably within real medication-safety workflows. The patient's narrative should not be preserved as decoration around structured reporting; it should remain an auditable source of evidence whose transformation is visible, limited, and open to correction.

ACKNOWLEDGMENTS: None
CONFLICT OF INTEREST: None
FINANCIAL SUPPORT: None
ETHICS STATEMENT: None

References

1. Rolfes L, van Hunsel F, van der Linden L, Taxis K, van Puijenbroek E. The quality of clinical information in adverse drug reaction reports by patients and healthcare professionals: A retrospective comparative analysis. Drug Saf. 2017;40(7):607-14. doi:10.1007/s40264-017-0530-5
2. O'Donovan B, Rodgers RM, Cox AR, Krska J. Identifying and managing adverse drug reactions: Qualitative analysis of patient reports to the UK Yellow Card Scheme. Br J Clin Pharmacol. 2022;88(7):3434-46. doi:10.1111/bcp.15263
3. Gordijn R, Wessels W, Kriek E, Nicolai MPJ, Elzevier HW, Visser L, et al. Patient reporting of sexual adverse events on an online platform for medication experiences. Br J Clin Pharmacol. 2022;88(12):5326-35. doi:10.1111/bcp.15454
4. Costa C, Abeijon P, Rodrigues DA, Figueiras A, Herdeiro MT, Torre C. Factors associated with underreporting of adverse drug reactions by patients: A systematic review. Int J Clin Pharm. 2023;45(6):1349-58. doi:10.1007/s11096-023-01592-y
5. Rolfes L, van Hunsel F, Caster O, Taavola H, Taxis K, van Puijenbroek E. Does patient reporting lead to earlier detection of drug safety signals? A retrospective comparison of time to reporting between patients and healthcare professionals in a global database. Br J Clin Pharmacol. 2018;84(7):1514-24. doi:10.1111/bcp.13576
6. Watson S, Chandler RE, Taavola H, Härmark L, Grundmark B, Zekarias A, et al. Safety concerns reported by patients identified in a collaborative signal detection workshop using VigiBase: Results and reflections from Lareb and Uppsala Monitoring Centre. Drug Saf. 2018;41(2):203-12. doi:10.1007/s40264-017-0594-2
7. de Vries ST, Denig P, Andrić A, Dimov Di Giusti M, Ptaszynska-Neophytou A, Härmark L, et al. Motives to report adverse drug reactions to the national agency: A survey study among healthcare professionals and patients in Croatia, The Netherlands, and the UK. Drug Saf. 2021;44(10):1073-83. doi:10.1007/s40264-021-01098-4
8. Robinson F, Wilkes S, Schaefer N, Goldstein M, Rice M, Gray J, et al. Patient-centered pharmacovigilance: Priority actions from the inherited bleeding disorders community. Ther Adv Drug Saf. 2023;14:20420986221146418. doi:10.1177/20420986221146418
9. Bate A, Hobbiger SF. Artificial intelligence, real-world automation and the safety of medicines. Drug Saf. 2021;44(2):125-32. doi:10.1007/s40264-020-01001-7
10. Edrees H, Song W, Syrowatka A, Simona A, Amato MG, Bates DW. Intelligent telehealth in pharmacovigilance: A future perspective. Drug Saf. 2022;45(5):449-58. doi:10.1007/s40264-022-01172-5
11. Leskur D, Bozic J, Rusic D, Seselja Perisin A, Cohadzic T, Pranic S, et al. Adverse drug reaction reporting via mobile applications: A narrative review. Int J Med Inform. 2022;168:104895. doi:10.1016/j.ijmedinf.2022.104895
12. Kassekert R, Grabowski N, Lorenz D, Schaffer C, Kempf D, Roy P, et al. Industry perspective on artificial intelligence/machine learning in pharmacovigilance. Drug Saf. 2022;45(5):439-48. doi:10.1007/s40264-022-01164-5
13. Laranjo L, Dunn AG, Tong HL, Kocaballi AB, Chen J, Bashir R, et al. Conversational agents in healthcare: A systematic review. J Am Med Inform Assoc. 2018;25(9):1248-58. doi:10.1093/jamia/ocy072
14. Kocaballi AB, Berkovsky S, Quiroz JC, Laranjo L, Tong HL, Rezazadegan D, et al. The personalization of conversational agents in health care: Systematic review. J Med Internet Res. 2019;21(11). doi:10.2196/15360
15. Tudor Car L, Dhinagaran DA, Kyaw BM, Kowatsch T, Joty S, Theng YL, et al. Conversational agents in health care: Scoping review and conceptual analysis. J Med Internet Res. 2020;22(8). doi:10.2196/17158
16. Martin GL, Jouganous J, Savidan R, Bellec A, Goehrs C, Benkebil M, et al. Validation of artificial intelligence to support the automatic coding of patient adverse drug reaction reports, using nationwide pharmacovigilance data. Drug Saf. 2022;45(5):535-48. doi:10.1007/s40264-022-01153-8
17. Létinier L, Jouganous J, Benkebil M, Bel-Létoile A, Goehrs C, Singier A, et al. Artificial intelligence for unstructured healthcare data: Application to coding of patient reporting of adverse drug reactions. Clin Pharmacol Ther. 2021;110(2):392-400. doi:10.1002/cpt.2266
18. Wasylewicz A, van de Burgt B, Weterings A, Jessurun N, Korsten E, Egberts T, et al. Identifying adverse drug reactions from free-text electronic hospital health record notes. Br J Clin Pharmacol. 2022;88(3):1235-45. doi:10.1111/bcp.15068
19. Davis SE, Zabotka L, Desai RJ, Wang SV, Maro JC, Coughlin K, et al. Use of electronic health record data for drug safety signal identification: A scoping review. Drug Saf. 2023;46(8):725-42. doi:10.1007/s40264-023-01325-0
20. Golder S, Xu D, O'Connor K, Wang Y, Batra M, Hernandez GG. Leveraging natural language processing and machine learning methods for adverse drug event detection in electronic health/medical records: A scoping review. Drug Saf. 2025;48(4):321-37. doi:10.1007/s40264-024-01505-6
21. Ding H, Simmich J, Vaezipour A, Andrews N, Russell T. Evaluation framework for conversational agents with artificial intelligence in health interventions: A systematic scoping review. J Am Med Inform Assoc. 2024;31(3):746-61. doi:10.1093/jamia/ocad222
22. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med. 2022;28(5):924-33. doi:10.1038/s41591-022-01772-9

23. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK, Chan AW, et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: The CONSORT-AI extension. Nat Med. 2020;26(9):1364-74. doi:10.1038/s41591-020-1034-x
24. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. 2024;385. doi:10.1136/bmj-2023-078378
25. Wiens J, Saria S, Sendak M, Ghassemi M, Liu VX, Doshi-Velez F, et al. Do no harm: A roadmap for responsible machine learning for health care. Nat Med. 2019;25(9):1337-40. doi:10.1038/s41591-019-0548-6
26. Morley J, Machado CCV, Burr C, Cowls J, Joshi I, Taddeo M, et al. The ethics of AI in health care: A mapping review. Soc Sci Med. 2020;260:113172. doi:10.1016/j.socscimed.2020.113172
27. Richardson JP, Smith C, Curtis S, Watson S, Zhu X, Barry B, et al. Patient apprehensions about the use of artificial intelligence in healthcare. NPJ Digit Med. 2021;4(1):140. doi:10.1038/s41746-021-00509-1
28. Seyyed-Kalantari L, Zhang H, McDermott MBA, Chen IY, Ghassemi M. Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations. Nat Med. 2021;27(12):2176-82. doi:10.1038/s41591-021-01595-0
29. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Med. 2019;17(1):195. doi:10.1186/s12916-019-1426-2
30. Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. Lancet Digit Health. 2021;3(11). doi:10.1016/S2589-7500(21)00208-9
31. Finlayson SG, Bowers JD, Ito J, Zittrain JL, Beam AL, Kohane IS. Adversarial attacks on medical machine learning. Science. 2019;363(6433):1287-9. doi:10.1126/science.aaw4399
32. Subbaswamy A, Saria S. From development to deployment: Dataset shift, causality, and shift-stable models in health AI. Biostatistics. 2020;21(2):345-52. doi:10.1093/biostatistics/kxz041